

This Page Is Inserted by IFW Operations
and is not a part of the Official Record

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images may include (but are not limited to):

- BLACK BORDERS
- TEXT CUT OFF AT TOP, BOTTOM OR SIDES
- FADED TEXT
- ILLEGIBLE TEXT
- SKEWED/SLANTED IMAGES
- COLORED PHOTOS
- BLACK OR VERY BLACK AND WHITE DARK PHOTOS
- GRAY SCALE DOCUMENTS

IMAGES ARE BEST AVAILABLE COPY.

**As rescanning documents *will not* correct images,
please do not report the images to the
Image Problem Mailbox.**

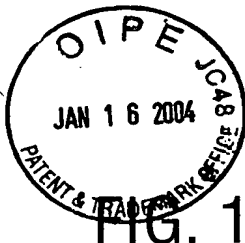
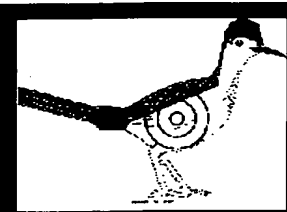


Fig. 1


1/40

File Edit View Go Communicator Help

Bookmarks Location http://Victoria.inpharmatica.co.uk/~volker/BPD3target.html



Target Mining Interface



Select Your Query Sequence

- Enter PDB accession number (e.g. 1QMA): and chain (e.g. B):
- OR
- Enter one Swiss-Prot accession (e.g. P27504) or GenBank proteinID (e.g. CAB08761.1):

Select Database

Release:

Apply Filters

- Iteration Filter: PSI-BLAST matches to be excluded:

If you select e.g. "Matches detected during the first 3 iterations" these matches will be excluded from the report (using the first_PB_iter annotation). This allows you to focus on more remote homologous which have been detected after 4 or more PSI-BLAST iterations. Matches detected using PSI-BLAST with negative iterations or using Genome-Threader are not effected by this option. However, if one match is found during the first e.g. 3 PSI-BLAST iterations and by Genome-Threader it will be excluded.

- Filter for the following SPECIES:
☐ Homo sapiens ☐ Rattus norvegicus (Rat) ☐ Mus musculus (Mouse) ☐ Danio rerio (Zebra fish)

100%

File Edit View Go Communicator Help

Bookmarks Location: http://London-bridge.inpharmatica.co.uk/cgi-bin/volker/getTargetBPD3.pl

2) 84 additional hits identified by both, Genome Threader and PSI-BLAST:

Combined Genome Threader and PSI - Blast output: PSI - BLAST values are shown in maroon!

Add2list	BPD links	WWW Link	Title	Organism	Div.	% ID (GT,FSI)	Query rqn. (GT,FSI)	Target rqn. (GT,FSI)	Aln. score (GT)	Conf. (GT)	1st Iter. (FSI)	Best Iter. (FSI)	Ex- v (FSI)
J	AAA59544.1 drill through Top50BlastHits Red Seq. View	AAA59544.1	Not given	Homo sapiens	PRI	33.7%, 33% <u>unmaskedSW</u>	4-183, 4-183	150-336, 150-336	449	100% <u>unmaskedGT</u>	1	2	3E-
→	AAB24821.1 drill through Top50BlastHits Red Seq. View	AAB24821.1	leukocyte integrin alpha chain	Homo sapiens	PRI	33.7%, 33% <u>unmaskedSW</u>	4-183, 4-183	150-336, 150-336	449	100% <u>unmaskedGT</u>	1	2	3E-
J	Q99715 drill through Top50BlastHits Red Seq. View	Q99715	COLLAGEN ALPHA 1(XII) CHAIN PRECURSOR.	Homo sapiens (Human).	PRI	28.8%, 26% <u>unmaskedSW</u>	4-181, 2-174	140-318, 2321-2495	440	100% <u>unmaskedGT</u>	1	3	2E-
J	P20701 drill through Top50BlastHits Red Seq. View	P20701	LEUKOCYTE ADHESION GLYCOPROTEIN LFA-1 ALPHA CHAIN PRECURSOR (LEUKOCYTE FUNCTION ASSOCIATED MOLECULE 1, ALPHA CHAIN)(CD11A) (INTEGRIN ALPHA-L).	Homo sapiens (Human).	PRI	100%, 100% <u>unmaskedSW</u>	1-183, 1-183	153-335, 153-335	423	100% <u>unmaskedGT</u>	1	1	1E-
J	AAC31672.1 drill through Top50BlastHits Red Seq. View	AAC31672.1	leukocyte function-associated molecule-1 alpha subunit	Homo sapiens	PRI	99.5%, 99% <u>unmaskedSW</u>	1-183, 1-183	153-335, 153-335	423	100% <u>unmaskedGT</u>	1	1	1E-
J	CAA72402.1 drill through Top50BlastHits Red Seq. View	CAA72402.1	collagen type XIV	Homo sapiens	PRI	29.1%, 29% <u>unmaskedSW</u>	2-180, 2-180	5-185, 5-185	422	100% <u>unmaskedGT</u>	1	2	2E-
J	AAB38702.1 drill through Top50BlastHits Red Seq. View	AAB38702.1	cartilage matrix protein	Homo sapiens	PRI	31.7%, 27% <u>unmaskedSW</u>	4-183, 2-182	275-455, 39-223	413	100% <u>unmaskedGT</u>	1	2	2E-
J	CAB70853.1 drill through Top50BlastHits Red Seq. View	CAB70853.1	hypothetical protein	Homo sapiens	PRI	28%, 28% <u>unmaskedSW</u>	1-180, 1-183	437-620, 437-624	406	100% <u>unmaskedGT</u>	1	2	1E-
J	CAA27972.1 drill through Top50BlastHits Red Seq. View	CAA27972.1	Not given	Homo sapiens	PRI	20.5%, 20% <u>unmaskedSW</u>	3-183, 2-181	1497-1873, 1689-1873	405	100% <u>unmaskedGT</u>	2	3	1E-
J	AAB59512.1 drill through Top50BlastHits Red Seq. View	AAB59512.1	Not given	Homo sapiens	PRI	20.5%, 20% <u>unmaskedSW</u>	3-183, 2-181	758-934, 950-1134	405	100% <u>unmaskedGT</u>	2	3	1E-
J	CAA07569.1 drill through Top50BlastHits Red Seq. View	CAA07569.1	matriin-4	Homo sapiens	PRI	28.1%, 25% <u>unmaskedSW</u>	1-183, 1-183	342-528, 31-217	403	100% <u>unmaskedGT</u>	1	2	2E-

FIG. 2B

1) 509 hits identified by Genome Threader only:											
Redundant sequence display	BPD link	WWW link	Title	Organism	Div.	%ID	Query rgn.	Target rgn.	Aln. score	Conf.	
Red Seq View	AAF71133.1 drill through Top50BlastHits	AAF71133.1	PR02769	Homo sapiens	PRI	13.8% unmaskedSW	108-179	1-80	122	100% unmaskedGI	
Red Seq View	CAB52192.1 drill through Top50BlastHits	CAB52192.1	G7c protein	Homo sapiens	PRI	9.6% unmaskedSW	7-117	20-124	82	99.18% unmaskedGI	
Red Seq View	CAA82910.1 drill through Top50BlastHits	CAA82910.1	basic transcription factor 2, 44 kD subunit	Homo sapiens	PRI	11% unmaskedSW	5-164	61-225	75	99.18% unmaskedGI	
Red Seq View	AAD21820.1 drill through Top50BlastHits	AAD21820.1	NG37	Homo sapiens	PRI	9.6% unmaskedSW	7-117	318-422	82	98.86% unmaskedGI	
Red Seq View	BAA20761.1 drill through Top50BlastHits	BAA20761.1	Not given	Homo sapiens	PRI	16.4% unmaskedSW	5-114	1836-1950	78	98.52% unmaskedGI	
Red Seq View	AAA36154.1 drill through Top50BlastHits	AAA36154.1	Not given	Homo sapiens	PRI	13.2% unmaskedSW	5-112	10-137	79	98.16% unmaskedGI	
Red Seq View	AAF03046.1 drill through Top50BlastHits	AAF03046.1	candidate tumor suppressor protein DICE1	Homo sapiens	PRI	13.8% unmaskedSW	5-113	4-131	79	97.07% unmaskedGI	
Red Seq View	AAC74854.1 drill through Top50BlastHits	AAC74854.1	orf, hypothetical protein	Escherichia coli	BCT	16.7% unmaskedSW	5-89	250-335	78	95.09% unmaskedGI	
Red Seq View	AAB60942.1 drill through Top50BlastHits	AAB60942.1	breast cancer suppressor candidate 1	Homo sapiens	PRI	16.2% unmaskedSW	63-156	2-80	80	93.33% unmaskedGI	
Red Seq View	AA67537.1 drill through Top50BlastHits	AA67537.1	glycoprotein IIIa	Homo sapiens	PRI	18.9% unmaskedSW	3-112	111-251	72	92.08% unmaskedGI reverse Hit	
Red Seq View	AA52589.1 drill through Top50BlastHits	AA52589.1	Not given	Homo sapiens	PRI	19.6% unmaskedSW	3-112	138-277	74	91.34% unmaskedGI reverse Hit	
Red Seq View	AAB71380.1 drill through Top50BlastHits	AAB71380.1	platelet membrane glycoprotein IIIa beta subunit	Homo sapiens	PRI	19.6% unmaskedSW	3-112	138-277	74	91.34% unmaskedGI reverse Hit	
Red Seq View	AAA35927.1 drill through Top50BlastHits	AAA35927.1	Not given	Homo sapiens	PRI	19.6% unmaskedSW	3-112	138-277	74	91.34% unmaskedGI	




6/40
FIG. 4


File Edit View Go Communicator Help

Back Forward Reload Home Search Netscape Print Security Stop

Bookmarks Location: <http://www.sanger.ac.uk/cgi-bin/Pfam/nph-search.cgi>



Pfam
Protein families database of alignments and HMMs
Home | Keyword search | Protein search | DNA search | Browse Pfam | Taxonomy search | Help

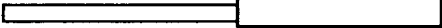


Results for gi|1788084|gb|AAC74854.1|

There were no matches to Pfam-A (including borderline matches) for gi|1788084|gb|AAC74854.1|

Matches to Pfam-B

Domain	Start	End	Evalue	Alignment
Pfam-B_39416	233	423	3.7e-103	Align

 [427 residues]

Alignments of Pfam-B domains to best-matching Pfam-B sequence

Format for fetching alignments to Pfam-B families ☐ Hypertext linked to swisspfam ☐

Query gi|1788084|gb|AAC74854.1|233-423 matching [Pfam-B_39416](#)

```

YEAH_ECOLI 233 DLRYKNYEKRDPDSSQAVMFCCLMDVSGSMDQSTKDMAKRFYILLYLFLSR 282
gi|1788084|gb|AAC74854.1| 233 DLRYKNYEKRDPDSSQAVMFCCLMDVSGSMDQSTKDMAKRFYILLYLFLSR 282

YEAH_ECOLI 283 TYKNVEVVYIRHDTQAKEVDEMEFFYSQETGGTIVSSALKLMDEVVKERY 332
gi|1788084|gb|AAC74854.1| 283 TYKNVEVVYIRHDTQAKEVDEMEFFYSQETGGTIVSSALKLMDEVVKERY 332

YEAH_ECOLI 333 NPAQWNIIYAAQASDGDNDWADD SPLCHEILAKKLLPVVRYYSYIEITRRAM 382
gi|1788084|gb|AAC74854.1| 333 NPAQWNIIYAAQASDGDNDWADD SPLCHEILAKKLLPVVRYYSYIEITRRAM 382

YEAH_ECOLI 383 QTLWREYEHQLSTFDNFAMQHIRDQDDIYPVVFRELPHKQNA 423
gi|1788084|gb|AAC74854.1| 383 QTLWREYEHQLSTFDNFAMQHIRDQDDIYPVVFRELPHKQNA 423

```

[Align to family](#)

If you think there is anything wrong with this script, please contact [Pfam](#)

100%

FIG. 5

7/40

File Edit View Go Communicator Help

Bookmarks Location: <http://www.ncbi.nlm.nih.gov:80/entrez/query.fcgi?cmd=Retrieve&db=>

LOCUS AAC74854 427 aa BCT 01-DEC-2000
 DEFINITION orf, hypothetical protein [Escherichia coli K12].
 ACCESSION AAC74854
 PID g1788084
 VERSION AAC74854.1 GI:1788084
 DBSOURCE locus AE000273 accession [AE000273.1](#)
 KEYWORDS
 SOURCE Escherichia coli K12.
 ORGANISM Escherichia coli K12
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae;
 Escherichia.
 REFERENCE 1 (residues 1 to 427)
 AUTHORS Blattner, F.R., Plunkett, G. III, Bloch, C.A., Perna, N.T., Burland, V.,
 Riley, M., Collado-Vides, J., Glasner, J.D., Rode, C.K., Mayhew, G.F.,
 Gregor, J., Davis, N.W., Kirkpatrick, M.A., Goeden, M.A., Rose, D.J.,
 Mau, B. and Shao, Y.
 TITLE The complete genome sequence of Escherichia coli K-12
 JOURNAL Science 277 (5331), 1453-1474 (1997)
 MEDLINE 97426617
 PUBMED 9278503
 REFERENCE 2 (residues 1 to 427)
 AUTHORS Blattner, F.R.
 TITLE Direct Submission
 JOURNAL Submitted (16-JAN-1997) Guy Plunkett III, Laboratory of Genetics,
 University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA.
 Email: ecoli@genetics.wisc.edu Phone: 608-262-2534 Fax:
 608-263-7459
 REFERENCE 3 (residues 1 to 427)
 AUTHORS Blattner, F.R.
 TITLE Direct Submission
 JOURNAL Submitted (02-SEP-1997) Guy Plunkett III, Laboratory of Genetics,
 University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA.
 Email: ecoli@genetics.wisc.edu Phone: 608-262-2534 Fax:
 608-263-7459
 REFERENCE 4 (residues 1 to 427)
 AUTHORS Plunkett, G. III.
 TITLE Direct Submission
 JOURNAL Submitted (13-OCT-1998) Laboratory of Genetics, University of
 Wisconsin, 445 Henry Mall, Madison, WI 53706, USA
 COMMENT This sequence was determined by the E. coli Genome Project at the
 University of Wisconsin-Madison (Frederick R. Blattner, director).
 Supported by NIH grants HG00301 and HG01428 (from the Human Genome
 Project and HCMGR). The entire sequence was independently
 determined from E. coli K12 strain M61655. Predicted open reading
 frames were determined using GenMark software, kindly supplied by
 Mark Borodovsky, Georgia Institute of Technology, Atlanta, GA,
 30332 [e-mail: mark@amber.gatech.edu]. Open reading frames that
 have been correlated with genetic loci are being annotated with CG
 Site Nos., unique ID nos. for the genes in the E. coli Genetic
 Stock Center (CGSC) database at Yale University, kindly supplied by
 Mary Berlyn. A public version of the database is accessible
 (<http://cgsc.biology.yale.edu>). Annotation of the genome is an
 ongoing task whose goal is to make the genome sequence more useful
 by correlating it with other data. Comments to the authors are
 appreciated. Updated information will be available at the E. coli
 Genome Project's World Wide Web site
 (<http://www.genetics.wisc.edu>). *** The E. coli K12 sequence and
 its annotations are periodically updated; this is version M54. No
 sequence changes. Annotation updates: updated gene identifications
 and products; all new functional assignments courtesy of Monica
 Riley; added promoters, protein binding sites, and repeated
 sequences described in reference 1. The unique numeric identifiers
 beginning with a lowercase 'b' assigned to each gene (protein- or
 RNA-encoding) are now designated as gene synonyms instead of
 labels. This should allow them to be searched for in Entrez as gene
 names.
 Method: conceptual translation.
 FEATURES
 source Location/Qualifiers
 1..427
 /organism="Escherichia coli K12"
 /strain="K12"
 /sub_strain="M61655"
 /db_xref="taxon:83333"
 Protein 1..427
 /function="orf; Unknown"
 /product="orf, hypothetical protein"
 CDS 1..427
 /gene="yeaH"
 /coded_by="1788078:6385..7668"
 /transl_table=11
 /note="o427. This 427 aa ORF is 28 pct identical (43 gaps)
 to 327 residues of an approx. 312 aa protein YZDC_BACSU
 SW: P45742"
 ORIGIN
 1 mtwfidrrln gknksmvnrq rflrrykaqi kqsiseaink rsvtdvdsge svsiptedis
 61 epmfhggrrg lhrvhpgnd hfvqndrier pggggggsgs gggqasqdg gqdefvfgis
 121 kdeyldllfe dialpnlkqn qqrqlteykt hrsgytangv panisvvrsl qnslarrtam
 181 taokrrrelha leenlaiisn sepaqlleee rlrkeiaelr akiervpfid tfdlryknye

100%

FIG. 6A

File Edit Search Type Help

Accession Code Query

inpharmatica

Tools

Sequences

Ligands

Select a database and enter a code:

Database	
GenBank	
Accession	<input checked="" type="radio"/>
Accession Version	<input type="radio"/>
GI-NID	<input type="radio"/>
protein_id	<input type="radio"/>
GI-PID	<input type="radio"/>
SWISS-PROT	
Accession	<input type="radio"/>
ID	<input type="radio"/>
PDB	
ID	<input type="radio"/>

AAC74854.1

Submit Reset

FIG. 6B

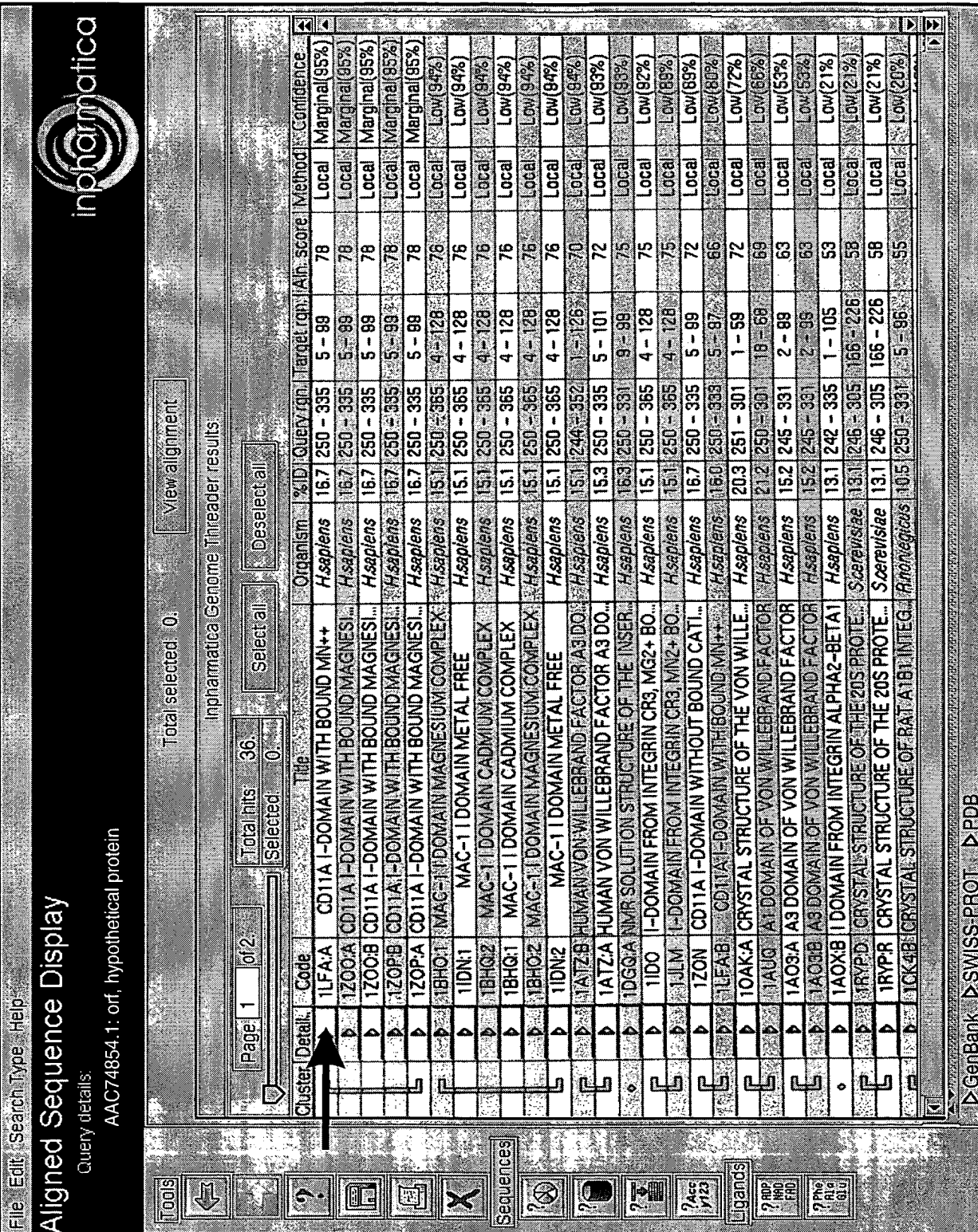


FIG. 6C

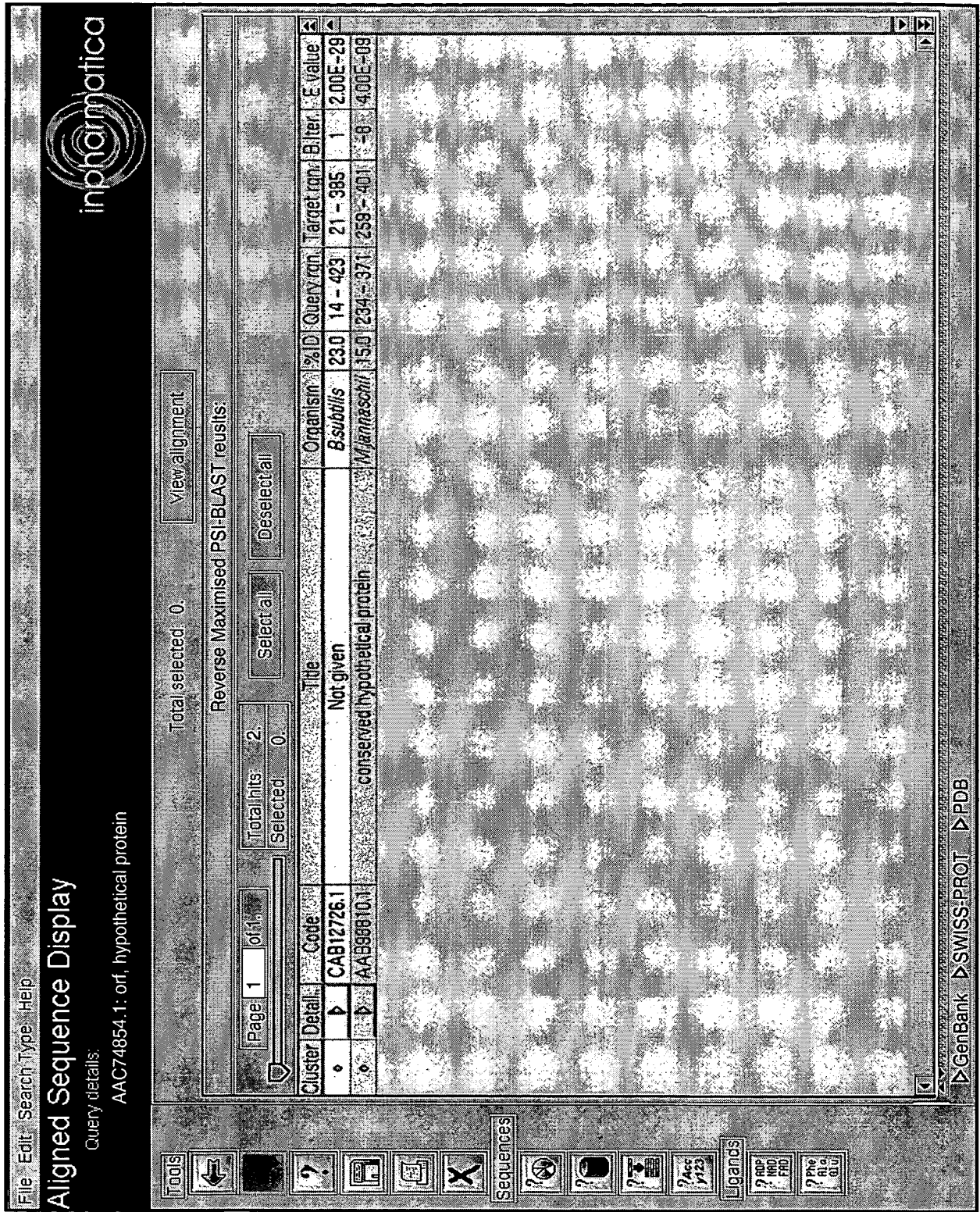


FIG. 7

AIEye output (January 4, 2002 3:07 PM)

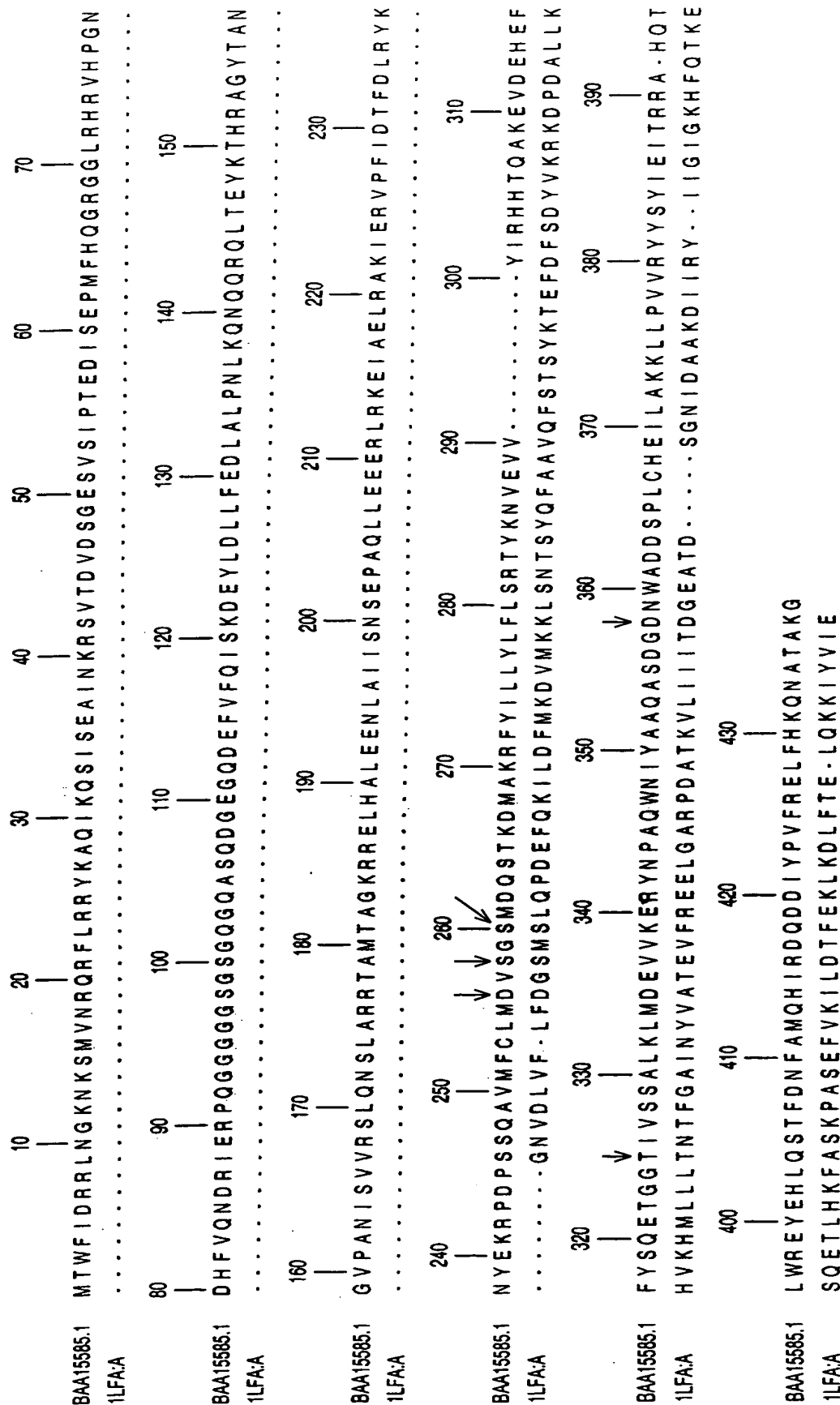
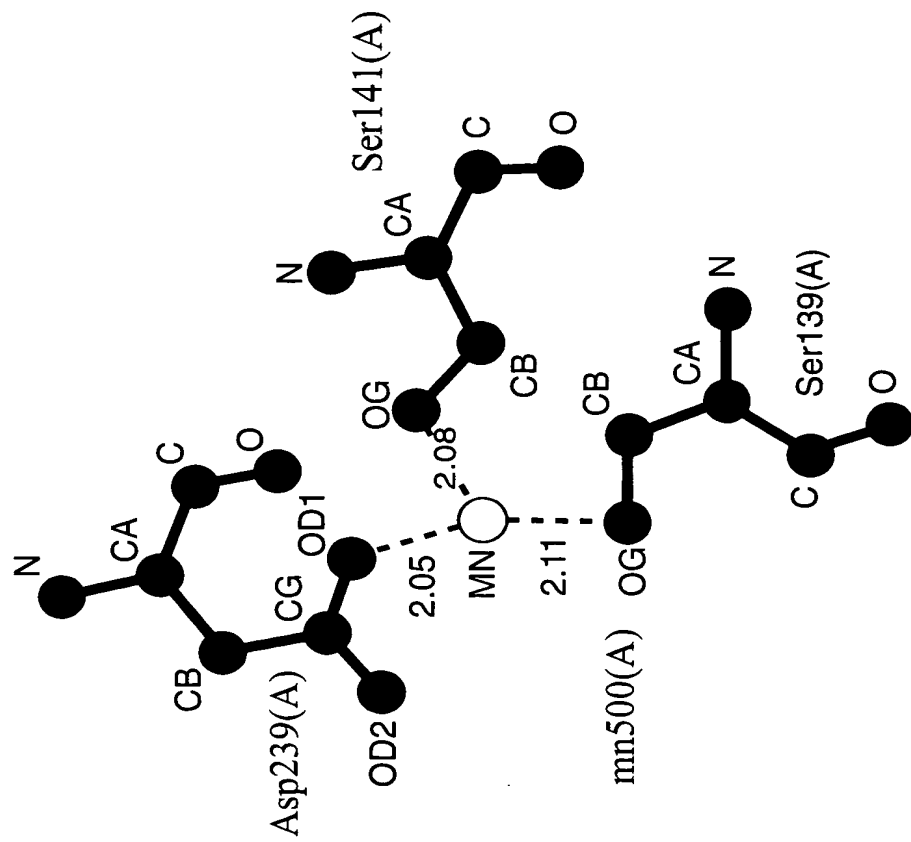


FIG. 8A



1lfa: MN500 Chain [A]

FIG. 8B

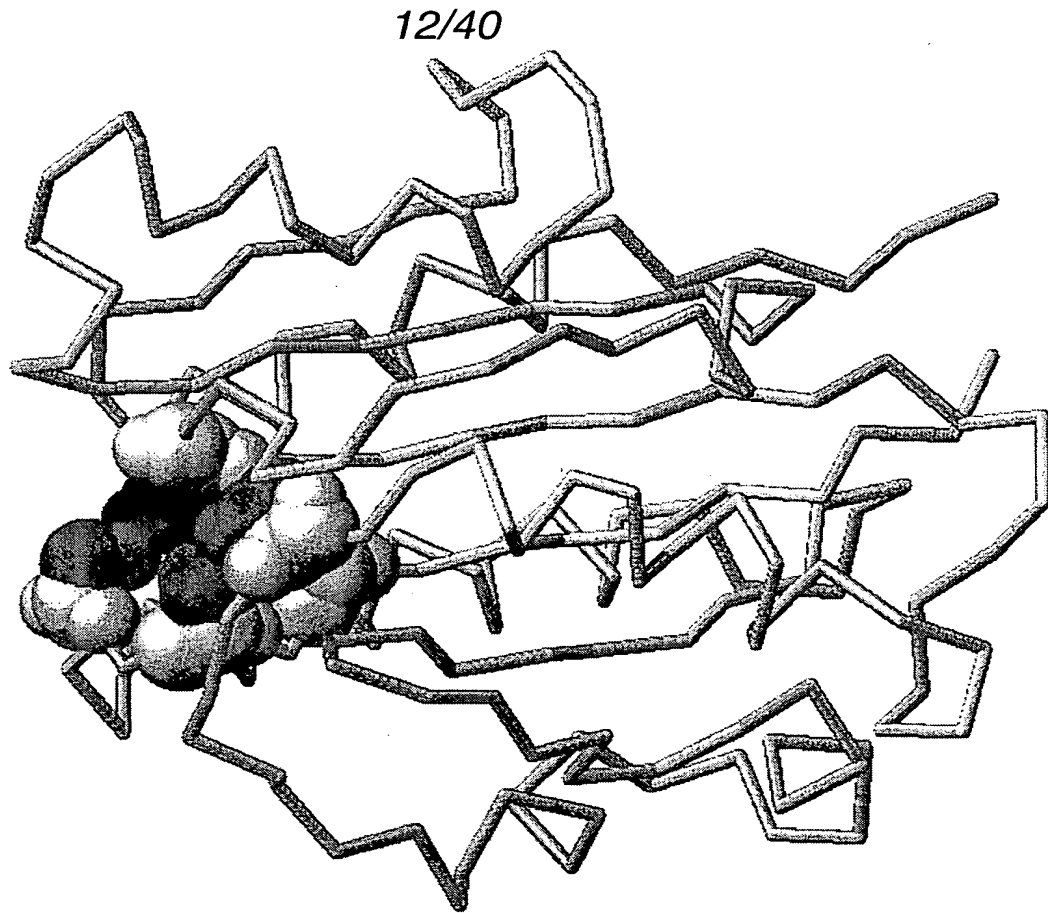
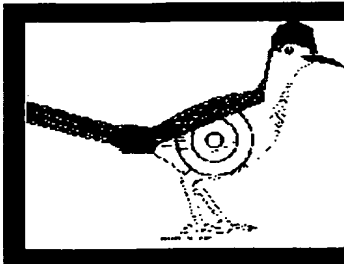



FIG. 9

13/40

File Edit View Go Communicator Help
Bookmarks Location http://victoria.inpharmatica.co.uk/~volker/BPD3target.html



Target Mining Interface



Select Your Query Sequence

- Enter PDB accession number (e.g. 1QMA): and chain (e.g. B):

OR

- Enter one Swiss-Prot accession (e.g. P27504) or GenBank proteinID (e.g. CAB08761.1):

Select Database

Release:

Apply Filters

- Iteration Filter: PSI-BLAST matches to be excluded:

If you select e.g. "Matches detected during the first 3 iterations" these matches will be excluded from the report (using the first_PB_iter annotation). This allows you to focus on more remote homologous which have been detected after 4 or more PSI-BLAST iterations. Matches detected using PSI-BLAST with negative iterations or using Genome-Threader are not effected by this option. However, if one match is found during the first e.g. 3 PSI-BLAST iterations and by Genome-Threader it will be excluded.

FIG. 10A

Accession	RefSeq Link	Title	Organism	Div	%ID (GI/F30)	Query len (GI/F30)	Target len (GI/F30)	Aln. score (GI)	Conf. (GI)	1st Iter. (F30)	Best Iter. (F30)	Best E-value (F30)
J	P56199 cd1 through I05508a1H1ts Ref Seq. View	INTEGRIN ALPHA-1 (LAMININ AND COLLAGEN RECEPTOR) (VLA-1) (CD49A)	Homo sapiens (Human)	PRI	51.8%, 51% unmasked SW	2-200, 2-200	139-337, 139-337	487	100% unmasked GI	1	2	2E-73
J	P17301 cd1 through I05508a1H1ts Ref Seq. View	PLATELET MEMBRANE GLYCOPROTEIN IA PRECURSOR (GPIA) (COLLAGEN RECEPTOR) (INTEGRIN ALPHA-2) (VLA-2 ALPHA CHAIN) (CD49B)	Homo sapiens (Human)	PRI	88.5%, 100% unmasked SW	2-200, 2-200	169-387, 169-387	471	100% unmasked GI	1	1	1E-111
J	AA01258.1 cd1 through I05508a1H1ts Ref Seq. View	Integrin alpha-11 subunit precursor	Homo sapiens	PRI	45.2%, 45% unmasked SW	2-200, 2-200	159-355, 159-355	469	100% unmasked GI	1	2	4E-68
J	AA05191.2 cd1 through I05508a1H1ts Ref Seq. View	Integrin alpha 11 subunit precursor	Homo sapiens	PRI	45.2%, 45% unmasked SW	2-200, 2-200	159-355, 159-355	469	100% unmasked GI	1	2	4E-68
J	Q99715 cd1 through I05508a1H1ts Ref Seq. View	COLLAGEN ALPHA (XII) CHAIN PRECURSOR.	Homo sapiens (Human)	PRI	28.1%, 31% unmasked SW	2-196, 7-196	133-318, 2323-2495	437	100% unmasked GI	1	2	2E-42
J	AA059491.1 cd1 through I05508a1H1ts Ref Seq. View	Not given	Homo sapiens	PRI	27.1%, 27% unmasked SW	2-200, 2-200	144-338, 144-338	435	100% unmasked GI	1	2	1E-60
J	AA059544.1 cd1 through I05508a1H1ts Ref Seq. View	Not given	Homo sapiens	PRI	27.1%, 27% unmasked SW	2-200, 2-200	144-338, 144-338	435	100% unmasked GI	1	2	1E-60
J	CA07122.1 cd1 through I05508a1H1ts Ref Seq. View	alpha2(BSD)15.1 (collagen, type XII, alpha 1)	Homo sapiens	PRI	26.8%, 26% unmasked SW	5-195, 156-201	293-472, 1591-1626	424	100% unmasked GI	2	3	5E-6
J	P21841 cd1 through I05508a1H1ts Ref Seq. View	CARTILAGE MATRIX PROTEIN PRECURSOR (MATRILIN-1)	Homo sapiens (Human)	PRI	26.8%, 26% unmasked SW	3-196, 3-196	271-455, 37-224	424	100% unmasked GI	1	2	3E-46
J	AA01506.1 cd1 through I05508a1H1ts Ref Seq. View	type XII collagen	Homo sapiens	PRI	27.6%, 30% unmasked SW	2-195, 2-195	133-318, 133-318	422	100% unmasked GI	1	2	3E-50
J	AA031952.1 cd1 through I05508a1H1ts Ref Seq. View	Integrin subunit alpha 10 precursor	Homo sapiens	PRI	46.5%, 46% unmasked SW	2-199, 2-199	162-359, 162-359	419	100% unmasked GI	1	2	5E-63
J	CA007569.1 cd1 through I05508a1H1ts	matrin-4	Homo sapiens	PRI	26%, 26% unmasked SW	1-199, 2-200	335-529, 28-223	411	100% unmasked GI	1	2	1E-47

2) 82 additional hits identified by both, Genome Threader and PSI-BLAST:

Combined Genome Threader and PSI-BLAST output: PSI-BLAST values are shown in mouse!

http://london-bridge.inpharmatica.co.uk/cgi-bin/volker/getTargetBPD3.pl#Mine2

FIG. 10B

GT
Confidence


	<u>Top50BlastHits</u> Red Seq View	<u>BAA91707.1</u>	Not given	<u>Human</u> sapiens	PRI	<u>unmasked SW</u> unmasked GT	<u>unmasked SW</u> unmasked GT	36-141	142	<u>unmasked GT</u>	3	5	2E-8
J	<u>AAA36795.1</u> drill through <u>Top50BlastHits</u>	<u>AAA36795.1</u>	undulin 2	Homo sapiens	PRI	34.7%, 21% unmasked SW	2-46, 139-196	3-51, 37-93	121	100% unmasked GT	2	2	8E-6
J	<u>AA40367.1</u> drill through <u>Top50BlastHits</u>	<u>AA40367.1</u>	calcium-activated chloride channel-2	Homo sapiens	PRI	15%, 15% unmasked SW	8-117, 8-117	312-421, 312-421	117	100% unmasked GT	5	5	1E-4
J	<u>CAA67559.1</u> drill through <u>Top50BlastHits</u>	<u>CAA67559.1</u>	collagen VI-alpha-1 chain	Homo sapiens	PRI	21.1%, 21% unmasked SW	108-197, 108-197	7-92, 7-92	114	100% unmasked GT	3	3	5E-7
→	<u>AAC76768.1</u> drill through <u>Top50BlastHits</u>	<u>AAC76768.1</u>	orf, hypothetical protein	Escherichia coli	BCT	13.2%, 11% unmasked SW	9-134, 7-186	267-384, 265-424	107	100% unmasked GT	-4	-5	3E-24
J	<u>CAB43000.1</u> drill through <u>Top50BlastHits</u>	<u>CAB43000.1</u>	tellurite resistance	Escherichia coli	BCT	12.9%, 14% unmasked SW	8-192, 2-181	214-398, 208-380	102	100% unmasked GT	-6	-7	6E-22
J	<u>AAA60114.1</u> drill through <u>Top50BlastHits</u>	<u>AAA60114.1</u>	platelet membrane glycoprotein IIb	Homo sapiens	PRI	14%, 13% unmasked SW	157-199, 157-199	145-187, 145-187	111	98.52% unmasked GT reverse Hit	-7	-8	8E-5



FIG. 12

File Edit View Go Communicator Help

Bookmarks Location: <http://www.sanger.ac.uk/cgi-bin/Pfam/nph-search.cgi>




The Sanger Centre

Pfam

Protein families database of alignments and HMMs

Home | Keyword search | Protein search | DNA search | Browse Pfam | Taxonomy search | Help




Results for gi|2367274|gb|AAC76768.1|

There were no matches to Pfam-A (including borderline matches) for gi|2367274|gb|AAC76768.1|

Matches to Pfam-B

Domain	Start	End	Evalue	Alignment
Pfam-B_15204	204	408	2.4e-108	Align


(427 residues)

Alignments of Pfam-B domains to best-matching Pfam-B sequence

Format for fetching alignments to Pfam-B families: Hypertext linked to swisspfam ☐

Query gi|2367274|gb|AAC76768.1|204-408 matching Pfam-B_15204

```

YIEM_ECOLI 204 DILRLLPPELATLGITELEYEFYRLVEKQLLTYRLNGESWREKVIERPV 253
gi|2367274|gb|AAC76768.1| 204 DILRLLPPELATLGITELEYEFYRLVEKQLLTYRLNGESWREKVIERPV 253

YIEM_ECOLI 254 VHKDYDEQPRGPFIVCVDTSGSHG6FNEQCAKAPCLALMRIALAENRRCY 303
gi|2367274|gb|AAC76768.1| 254 VHKDYDEQPRGPFIVCVDTSGSHG6FNEQCAKAPCLALMRIALAENRRCY 303

YIEM_ECOLI 304 IMLPSTEIVRYELSGPQGIEQAIRFLSQQFRG6TDLASCFRATMERLQSR 353
gi|2367274|gb|AAC76768.1| 304 IMLPSTEIVRYELSGPQGIEQAIRFLSQQFRG6TDLASCFRATMERLQSR 353

YIEM_ECOLI 354 EWPADADRVVISDFIAQRLPDDVT SKVKELQRVHQHMFHVAHMSAN6KP6I 403
gi|2367274|gb|AAC76768.1| 354 EWPADADRVVISDFIAQRLPDDVT SKVKELQRVHQHMFHVAHMSAN6KP6I 403

YIEM_ECOLI 404 MRIFD 408
gi|2367274|gb|AAC76768.1| 404 MRIFD 408
  
```

Align to family

If you think there is anything wrong with this script, please contact [Pfam](#)

File Edit View Go Communicator Help

Bookmarks Location: <http://www.ncbi.nlm.nih.gov:80/entrez/query.fcgi?cmd=Retrieve&db=protein>

LOCUS AAC76768 427 aa BCT 01-DEC-2000
 DEFINITION orf, hypothetical protein [Escherichia coli K12].
 ACCESSION AAC76768
 PID g2367274
 VERSION AAC76768.1 GI: 2367274
 DBSOURCE locus AE000451 accession [AE000451.1](#)
 KEYWORDS
 SOURCE Escherichia coli K12.
 ORGANISM Escherichia coli K12
 Bacteria; Proteobacteria; gamma subdivision; Enterobacteriaceae; Escherichia.

REFERENCE 1 (residues 1 to 427)
 AUTHORS Blattner, F.R., Plunkett, G. III, Bloch, C.A., Perna, R.T., Burland, V., Riley, M., Collado-Vides, J., Glasner, J.D., Rode, C.K., Mayhew, G.F., Gregor, J., Davis, N.W., Kirkpatrick, M.A., Goeden, M.A., Rose, D.J., Mau, B. and Shao, Y.
 TITLE The complete genome sequence of Escherichia coli K-12
 JOURNAL Science 277 (5331), 1453-1474 (1997)
 MEDLINE 97426612
 PUBMED 9278503

REFERENCE 2 (residues 1 to 427)
 AUTHORS Blattner, F.R.
 TITLE Direct Submission
 JOURNAL Submitted (16-JAN-1997) Guy Plunkett III, Laboratory of Genetics, University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA. Email: ecoli@genetics.wisc.edu Phone: 608-262-2534 Fax: 608-263-7459

REFERENCE 3 (residues 1 to 427)
 AUTHORS Blattner, F.R.
 TITLE Direct Submission
 JOURNAL Submitted (02-SEP-1997) Guy Plunkett III, Laboratory of Genetics, University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA. Email: ecoli@genetics.wisc.edu Phone: 608-262-2534 Fax: 608-263-7459

REFERENCE 4 (residues 1 to 427)
 AUTHORS Plunkett, G. III.
 TITLE Direct Submission
 JOURNAL Submitted (13-OCT-1998) Laboratory of Genetics, University of Wisconsin, 445 Henry Mall, Madison, WI 53706, USA

COMMENT This sequence was determined by the E. coli Genome Project at the University of Wisconsin-Madison (Frederick R. Blattner, director). Supported by NIH grants M600301 and M601428 (from the Human Genome Project and HCHGR). The entire sequence was independently determined from E. coli K12 strain M61655. Predicted open reading frames were determined using Genemark software, kindly supplied by Mark Borodovsky, Georgia Institute of Technology, Atlanta, GA, 30332 [e-mail: mark@amber.gatech.edu]. Open reading frames that have been correlated with genetic loci are being annotated with CG Site Nos., unique ID nos. for the genes in the E. coli Genetic Stock Center (CGSC) database at Yale University, kindly supplied by Mary Berlyn. A public version of the database is accessible (<http://cgsc.biology.yale.edu>). Annotation of the genome is an ongoing task whose goal is to make the genome sequence more useful by correlating it with other data. Comments to the authors are appreciated. Updated information will be available at the E. coli Genome Project's World Wide Web site (<http://www.genetics.wisc.edu>). *** The E. coli K12 sequence and its annotations are periodically updated; this is version M54. No sequence changes. Annotation updates: updated gene identifications and products; all new functional assignments courtesy of Monica Riley; added promoters, protein binding sites, and repeated sequences described in reference 1. The unique numeric identifiers beginning with a lowercase 'b' assigned to each gene (protein- or RNA-encoding) are now designated as gene synonyms instead of labels. This should allow them to be searched for in Entrez as gene names.
 Method: conceptual translation.

FEATURES
 source Location/qualifiers
 1..427
 /organism="Escherichia coli K12"
 /strain="K12"
 /sub_strain="M61655"
 /db_xref="taxon:83333"
 Protein 1..427
 /function="orf, Unknown"
 /product="orf, hypothetical protein"
 CDS 1..427
 /gene="pieD"
 /coded_by="complement(2367272:5249..6532)"
 /transl_table=11
 /note="f427; sequence change joins ORFs pieD and pieM from earlier version"

ORIGIN
 1 mrsrlkdarv ppeteevnc yqqsqllstp qfivqlpqil dlhrlnspw aeqarqlvda
 61 nstitsalht lflgrwrlsl ivgattlmgq lleeereqll sevgermtls gglepiladn
 121 ntaagrldwm sagqlkrqdy qlivypgefl neqpalrla eqlgrsreak siprndagme
 181 tfrtmvzpa tvpeqvdlq qddilrlp delatloite leuefnrrls ebllltmrth

100%

FIG. 14A

The screenshot displays the 'inpharmatica' web application interface for an 'Accession Code Query'. The top navigation bar includes 'File', 'Edit', 'Search', 'Type', and 'Help'. The main content area is titled 'Accession Code Query' and features a sidebar on the left with icons for 'Tools', 'Sequences', and 'Ligands'. The 'Tools' section is currently active, showing a list of databases: 'GenBank', 'SWISS-PROT', and 'PDB'. Each database has a list of accession codes with radio buttons for selection. The 'GenBank' database is selected, and the 'Accession' field is highlighted. The 'Accession' field contains the value 'AAC76768.1'. Below the database list, there are 'Submit' and 'Reset' buttons. The 'Sequences' section is also visible, showing a list of sequences with a '?' icon next to each. The 'Ligands' section is partially visible at the bottom, showing a list of ligands with a 'Phe' icon next to each.

File Edit Search Type Help

Accession Code Query

inpharmatica

Tools

Sequences

Ligands

Select a database and enter a code:

Database	Accession
GenBank	Accession
	Accession Version
	GI-NUD
	protein_id
SWISS-PROT	Accession
	ID
PDB	ID

AAC76768.1

Submit Reset

AAC76768.1:orf, hypothetical protein

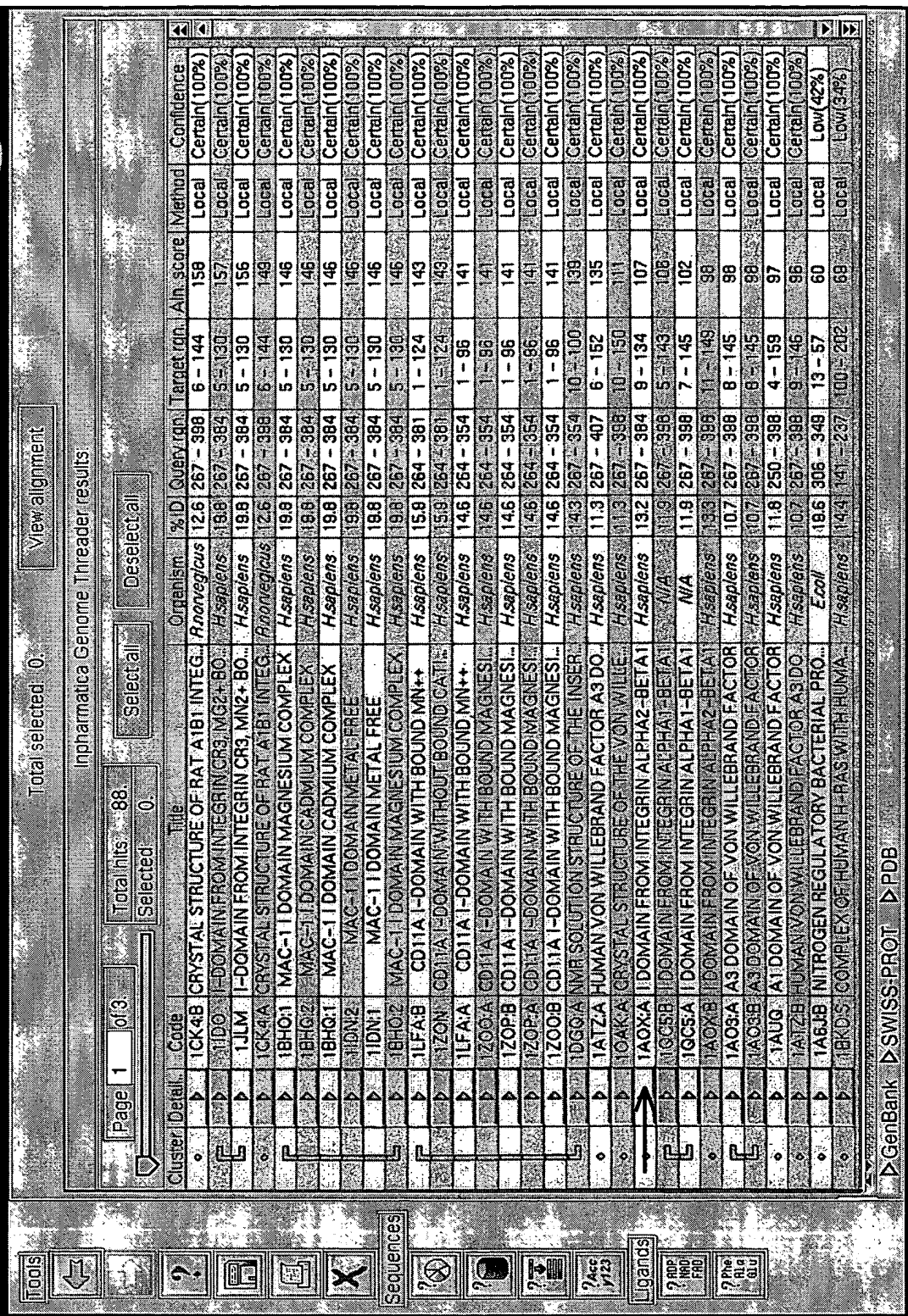


FIG. 15

AIEye output (December 13, 2000 3:07 PM)

```

AAC76768.1      10      20      30      40      50      60      70
1A0X:A      |      |      |      |      |      |      |
      M P S R L K D A R V P P E L T E E V M C Y Q Q S Q L L S T P Q F I V Q L P Q I D L L H R L N S P W A E Q A R Q L V D A N S T I T S A L H T L F L Q R W R
-----
AAC76768.1      80      90      100      110      120      130      140      150
1A0X:A      |      |      |      |      |      |      |
      L S L I V Q A T T L N Q Q L L E E R E R E Q L L S E V Q E R M T L S G Q L E P I L A D N N T A A G R L W D M S A G Q L K R G D Y Q L I V K Y G E F L N E Q P
-----
AAC76768.1      160      170      180      190      200      210      220      230
1A0X:A      |      |      |      |      |      |      |
      E L K R L A E Q L G R S R E A K S I P R N D A Q M E T F R T M V R E P A T V P E Q V D G L Q Q S D D I L R L P P E L A T L G I T E L E Y E F Y R R L V E
-----
AAC76768.1      240      250      260      270      280      290      300
1A0X:A      |      |      |      |      |      |      |
      K Q L L T Y R L H G E S W R E K V I E R P V V H K D Y D E Q P R G P F I V C V D T S G S M G G F N E Q C A K A F C L A L M R I -- A L A E N R R C Y I M L
-----
AAC76768.1      310      320      330      340      350      360      370      380
1A0X:A      |      |      |      |      |      |      |
      -- S C P S L I D -- V V V C D E S N S I Y P W D A -- V K N F L E K F V Q G L D I G P T K T Q V G L I Q
      F S T E I - V R Y E L S G P Q G -- I E Q A I R F L S Q Q F R G G T D L A S C F R A I M E R L Q S -- -- R E W F D A D A V V I S D F I A Q R L P D D
      Y A N N P R V V F N L N T Y K T K E E M I V A T S Q T S Q Y G G D L N T F G A I Q Y A R K Y A Y S A A S G G R R S A T K Y M V V V T D G E S H - D G S M
-----
AAC76768.1      390      400      410      420      430      440      450      460
1A0X:A      |      |      |      |      |      |      |
      V T S K V K E L Q R V H Q H R F H A V A M S A H G K P G I M R I F D H I W R F D T G M R S R L L R R W R P
      L K A V I D Q C N H D N - I L R F G I A V L G Y - - - - - L N R N A L D T K N L I K E I K A I A S I P T E R Y F F N V S D E A A L L E K A G
-----
AAC76768.1      470
1A0X:A      |
      T L G E Q I F S I E G G T

```


FIG. 16B

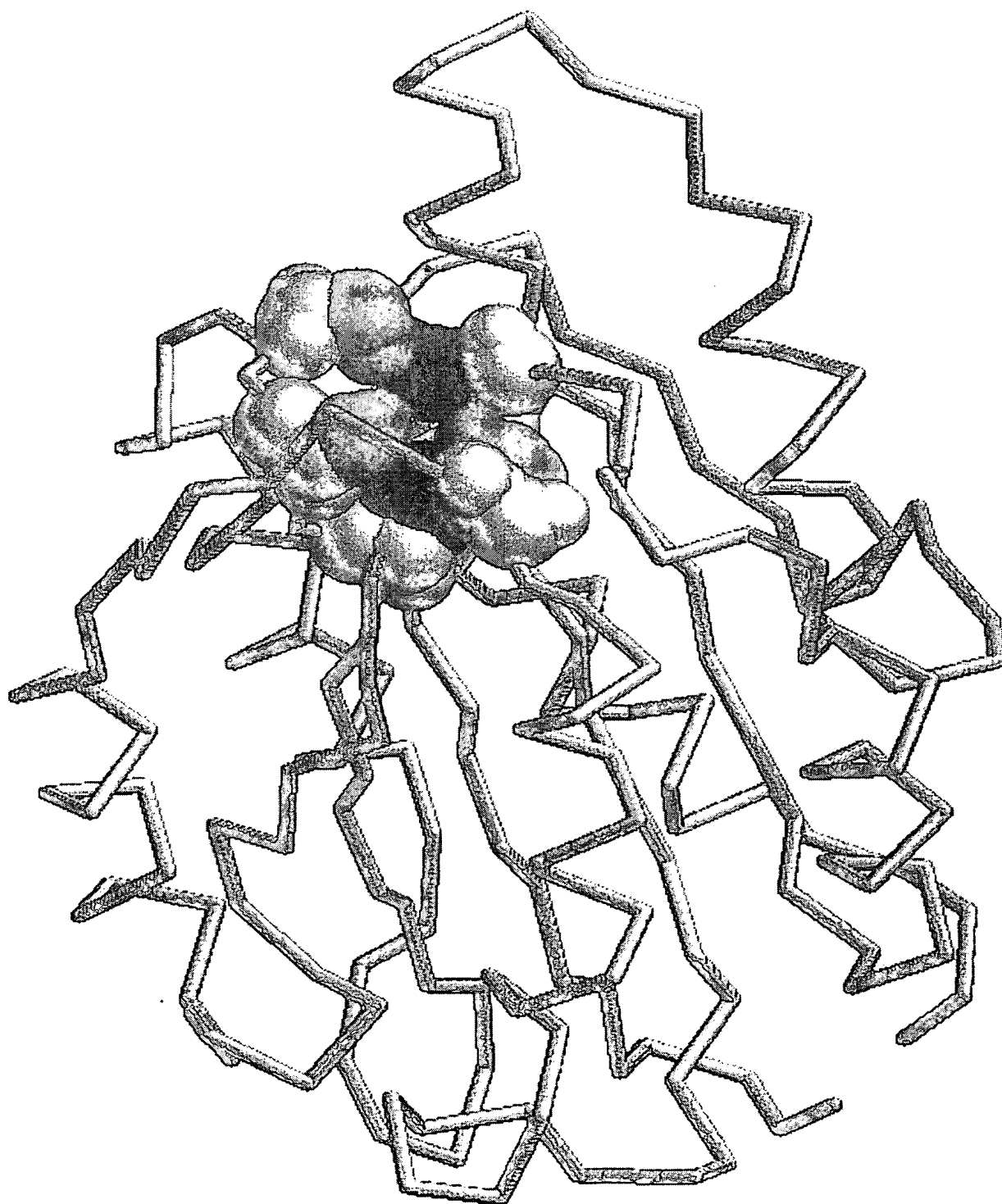
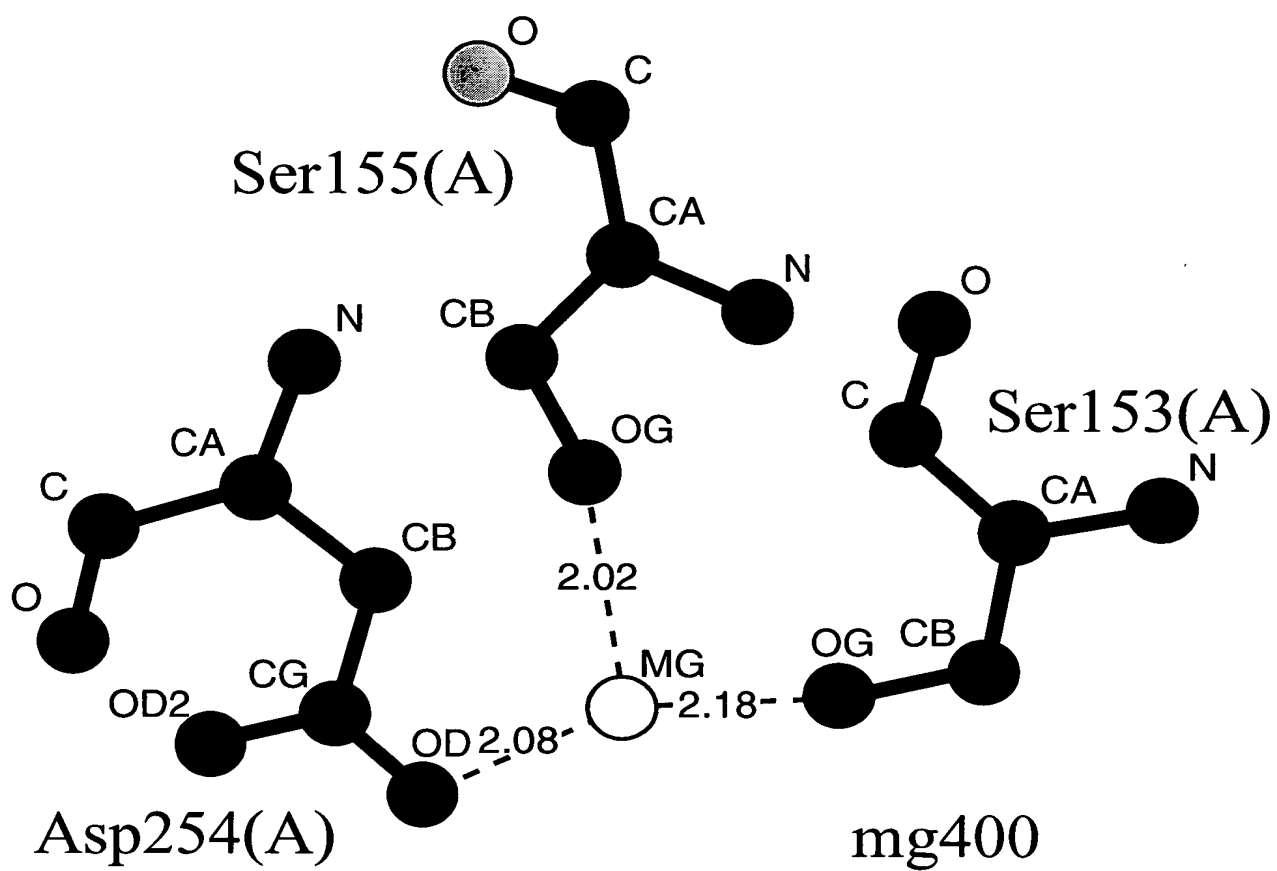
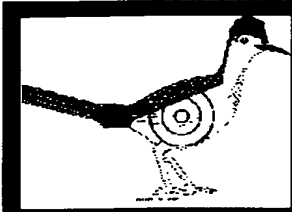



FIG. 16A



1aox: MG400

File Edit View Go Communicator Help
Bookmarks Location http://victoria.inpharmatica.co.uk/~volker/BPD3target.html

**Target Mining Interface**

inpharmatica

Select Your Query Sequence

- Enter PDB accession number (e.g. 1QMA): and chain (e.g. B):
- OR
- Enter one Swiss-Prot accession (e.g. P27504) or GenBank proteinID (e.g. CAB08761.1):

Select Database

Release:

Apply Filters

- Iteration Filter: PSI-BLAST matches to be excluded:

If you select e.g. "Matches detected during the first 3 iterations" these matches will be excluded from the report (using the first_PB_iter annotation). This allows you to focus on more remote homologous which have been detected after 4 or more PSI-BLAST iterations. Matches detected using PSI-BLAST with negative iterations or using Genome-Threader are not effected by this option. However, if one match is found during the first e.g. 3 PSI-BLAST iterations and by Genome-Threader it will be excluded.

- Filter for the following SPECIES:
☐ Homo sapiens ☐ Rattus norvegicus (Rat) ☐ Mus musculus (Mouse) ☐ Danio rerio (Zebra fish)

100%

File Edit View Go Communicator Help

Back Forward Reload Home Search Netscape Print Security Stop

Bookmarks Location: http://London-bridge.inpharmatica.co.uk/cgi-bin/volker/getTargetBPD3.pl

2) 81 additional hits identified by both, Genome Threader and PSI-BLAST:

Combined Genome Threader and PSI - Blast output: PSI - BLAST values are shown in maroon!

Add2list	BPD links	WWW link	Title	Organism	Div.	%ID (GT,PSI)	Query rgn. (GT,PSI)	Target rgn. (GT,PSI)	Aln. score (GT)	Conf. (GT)	1st Iter. (PSI)
J	AA459544.1 drill through Top50BlastHits Red Seq View	AA459544.1	Not given	Homo sapiens	PRI	100%, 100% unmaskedSW	1-187, 1-187	148-334, 148-334	488	100% unmaskedGT	1
→	AAB24821.1 drill through Top50BlastHits Red Seq View	AAB24821.1	leukocyte integrin alpha chain	Homo sapiens	PRI	100%, 100% unmaskedSW	1-187, 1-187	148-334, 148-334	488	100% unmaskedGT	1
J	Q99715 drill through Top50BlastHits Red Seq View	Q99715	COLLAGEN ALPHA 1(XII) CHAIN PRECURSOR.	Homo sapiens (Human)	PRI	28.9%, 28% unmaskedSW	2-186, 2-179	439-617, 2322-2494	456	100% unmaskedGT	1
J	AAB38702.1 drill through Top50BlastHits Red Seq View	AAB38702.1	cartilage matrix protein	Homo sapiens	PRI	28.9%, 25% unmaskedSW	2-186, 2-186	274-452, 40-221	446	100% unmaskedGT	1
J	AAC01506.1 drill through Top50BlastHits Red Seq View	AAC01506.1	type XII collagen	Homo sapiens	PRI	28.4%, 28% unmaskedSW	2-186, 2-186	137-318, 137-318	445	100% unmaskedGT	1
J	CAA72402.1 drill through Top50BlastHits Red Seq View	CAA72402.1	collagen type XIV	Homo sapiens	PRI	28.7%, 30% unmaskedSW	2-186, 2-186	6-185, 6-185	442	100% unmaskedGT	1
J	AAB38547.1 drill through Top50BlastHits Red Seq View	AAB38547.1	leukointegrin alpha d chain	Homo sapiens	PRI	61%, 60% unmaskedSW	1-187, 1-187	148-334, 148-334	439	100% unmaskedGT	1
J	CAB71222.1 drill through Top50BlastHits Red Seq View	CAB71222.1	dJ238D15.1 (collagen, type XII, alpha 1)	Homo sapiens	PRI	27.1%, 22% unmaskedSW	1-186, 2-186	293-472, 1430-1620	439	100% unmaskedGT	1
J	CAA07569.1 drill through Top50BlastHits Red Seq View	CAA07569.1	matrilin-4	Homo sapiens	PRI	27.9%, 24% unmaskedSW	2-186, 3-187	344-525, 34-215	418	100% unmaskedGT	2
J	CAB46380.1 drill through Top50BlastHits Red Seq View	CAB46380.1	dJ453C12.3 (matrilin-4)	Homo sapiens	PRI	27.9%, 24% unmaskedSW	2-186, 3-186	385-566, 34-214	418	100% unmaskedGT	2

FIG. 18B

File Edit View Go Communicator Back Forward Reload Home Search News Print Security Shop Help														
Bookmarks Location http://london-bridge.inpharmatica.co.uk/cgi-bin/volker/getTargetBPD3.pl														
Red Sea View	drill through	BAA31811.L	KIA0036 protein					Homo sapiens	PRI	67.2%	73-105	330-503	66	unmasked reverse hit
Red Sea View	drill through	CAB52192.L	G7c protein					Homo sapiens	PRI	14.1%	6-88	20-105	76	89.66% unmasked
Red Sea View	drill through	BAA75889.L	N-copine					Homo sapiens	PRI	28.8%	73-143	405-474	73	89.66% unmasked
Red Sea View	drill through	P26012	INTEGRIN BETA-8 PRECURSOR					Homo sapiens (human)	PRI	13.1%	2-115	147-289	78	87.73% unmasked
Red Sea View	drill through	AA021820.L	NG37					Homo sapiens	PRI	14.1%	6-88	318-403	76	85.59% unmasked
Red Sea View	drill through	CAA65775.L	Sec23 protein					Homo sapiens	PRI	19.7%	7-141	134-275	74	84.45% unmasked
Red Sea View	drill through	AAC15920.L	copine I					Homo sapiens	PRI	20%	73-185	384-497	67	80.79% unmasked reverse hit
Red Sea View	drill through	AAC63280.L	polymerase					Homo sapiens	PRI	19.8%	84-185	442-552	55	74.2% unmasked
Red Sea View	drill through	CAB87610.L	cJ009K20.3 (Copine (similar to KIA0036))					Homo sapiens	PRI	25%	76-143	363-429	67	72.84% unmasked
Red Sea View	drill through	AAA35532.L	Not given					Homo sapiens	PRI	11.1%	6-145	373-504	56	71.48% unmasked
Red Sea View	drill through	P10155	60 KDA RO PROTEIN (60 KDA RIBONUCLEOPROTEIN P01 (PORN1) (SUGREN SYNDROME TYPE A ANTIGEN (SS-A)))					Homo sapiens (human)	PRI	11.1%	6-145	373-504	56	70.11% unmasked
Red Sea View	drill through	AAC15863.L	hexokinase I					Homo sapiens	PRI	18.5%	22-122	749-861	75	68.76% unmasked reverse hit
Red Sea View	drill through	BAA92872.L	KIAA1434 protein					Homo sapiens	PRI	20.8%	124-186	576-647	67	67.41% unmasked
Red Sea View	drill through	CAA10335.L	Sec24B protein					Homo sapiens	PRI	9%	6-91	682-801	61	67.41% unmasked
Red Sea View	drill through	AA043766.L	prolactin gamma A11 short form protein					Homo sapiens	PRI	15.4%	30-185	475-643	66	66.06% unmasked
Red Sea View	drill through	AA043714.L	prolactin gamma A11					Homo sapiens	PRI	15.4%	30-185	475-643	66	63.4% unmasked
Red Sea View	drill through	CAA08933.L	annexin 31 (annexin XXXI)					Homo sapiens	PRI	10.2%	20-184	123-308	58	63.4% unmasked
Red Sea View	drill through	AAAS2646.L	hexokinase 1					Homo sapiens	PRI	18.5%	22-122	750-852	73	62.08% unmasked reverse hit
Red Sea View	drill through	AAAS9185.L	integrin beta-7 subunit					Homo sapiens	PRI	18.4%	2-55	152-204	67	59.48% unmasked
Red Sea View	drill through	AA85421.L	integrin variant betaE					Homo sapiens	PRI	24.3%	2-36	130-166	73	58.92% unmasked

FIG. 19

Biopendium additional annotation page - Netscape

File Edit View Go Communicator Help

Back Forward Reload Home Search Netscape Print Security Stop

Bookmarks Location: /london-bridge/inpharmatica/BPDDEV7/cgi-bin/isp.pl?rep_esn=55817&password=carss_app&username=carss_app&oracle_sid=BPDDEV7 What's Related

Aligned annotation view for P10155 (downloading image...)

1: AAA35450.1
Rep: P10155

1 50 100 150 200 250 300 350 400 450 500

Primary database information: ☒ SwissProt features

Secondary database information: ☒ Prosite-matches ☒ Prints matches

Inpharmatica calculated information: ☒ Secondary structure ☒ Ligand binding residues ☒ Masked regions

Sequence information

Source databank	SWISSPROT
Accession code	P10155
Gene name	SSA2 OR R060
Download sequence in FASTA format	

Links to other resources:

View custom hyperlinks

Search in ExPASy

European Bioinformatics Institute	IPR000504
Online Mendelian Inheritance in Man	234700
Online Mendelian Inheritance in Man	600063
Protein Information Resource	A30596
Protein Information Resource	A31760

Swissprot comments:

FUNCTION

UNKNOWN

RO SMALL RIBONUCLEOPROTEINS CONSIST OF FOUR SMALL RNA MOLECULES OF 85-112 NT, EACH OF

Document Done

31/40

FIG. 20

File Edit View Go Communicator

Back Forward Reload Home Search Netscape Print Security Stop

Bookmarks Location: http://www.sanger.ac.uk/cgi-bin/Pfam/nph-search.cgi

Pfam

Protein families database of alignments and HMMs

Home | Keyword search | Protein search | DNA search | Browse Pfam | Taxonomy search | Help

Results for gi|133251|sp|P10155|RO60_HUMAN

There were no matches to Pfam-A (including borderline matches) for gi|133251|sp|P10155|RO60_HUMAN

Matches to Pfam-B

Domain	Start	End	Evalue	Alignment
Pfam-B_8344	1	194	2.3e-103	Align
Pfam-B_10162	195	538	1.8e-165	Align

[538 residues]

Alignments of Pfam-B domains to best-matching Pfam-B sequence

Format for fetching alignments to Pfam-B families: ☒ Hypertext linked to swisspfam ☐

Query gi|133251|sp|P10155|RO60_HUMAN/1-194 matching Pfam-B_8344

```

q92787 1 MEE SVNQMQLNEKQIANSDGCVVQVTDMMRLMRFLCFGSEGGTTYIKE 50
gi|133251|sp|P10155|RO60_HUMAN 1 MEE SVNQMQLNEKQIANSDGCVVQVTDMMRLMRFLCFGSEGGTTYIKE 50

q92787 51 QKLGLENAEALIRLIEDGRGCEVIQEIKSFSEGRITTKQEPMLFALAI CS 100
gi|133251|sp|P10155|RO60_HUMAN 51 QKLGLENAEALIRLIEDGRGCEVIQEIKSFSEGRITTKQEPMLFALAI CS 100

q92787 101 QCSDISTKQAPFAVSEVCRIPTHLFTPIQFKKDLKESMKCGMWGRALRK 150
gi|133251|sp|P10155|RO60_HUMAN 101 QCSDISTKQAPFAVSEVCRIPTHLFTPIQFKKDLKESMKCGMWGRALRK 150

q92787 151 AIADWYNEKGGMALALAVTKYKQRNGWSHKDLLRLSHLKPSSSE 194
gi|133251|sp|P10155|RO60_HUMAN 151 AIADWYNEKGGMALALAVTKYKQRNGWSHKDLLRLSHLKPSSSE 194

```

Align to family

Query gi|133251|sp|P10155|RO60_HUMAN/195-538 matching Pfam-B_10162

```

008848 195 LAIVTKYITKGKEVMEYKEKAL SVEAEKLLKYLEAVEKVKRTKDDLEV 244
gi|133251|sp|P10155|RO60_HUMAN 195 LAIVTKYITKGKEVMEYKEKAL SVEAEKLLKYLEAVEKVKRTKDDLEV 244

008848 245 IMLEENQLVREMLLTNHLKSEVWKALLQEMPLTALLRNLGKMTANSVL 294
gi|133251|sp|P10155|RO60_HUMAN 245 IMLEENQLVREMLLTNHLKSEVWKALLQEMPLTALLRNLGKMTANSVL 294

008848 295 EPGNSEVSLICEKLSHEKLLKKARINPFFMLIALETYRAGHGLRGLKVI 344
gi|133251|sp|P10155|RO60_HUMAN 295 EPGNSEVSLICEKLSHEKLLKKARINPFFMLIALETYRAGHGLRGLKVI 344

008848 345 PDKDILQLDRAFYTTFTKVEPTGKRFLLAVDVSAEMHQRALGSVLNAST 394

```

32/40
FIG. 21

File Edit View Go Communicator Help

Back Forward Reload Home Search Netscape Print Security Stop

Bookmarks Location: <http://www.ncbi.nlm.nih.gov:80/entrez/query.fcgi?cmd=Retrieve&db=Prot>

LOCUS R060_HUMAN 538 aa PRI 01-FEB-1996
DEFINITION 60 KD RO PROTEIN (60 KD RIBONUCLEOPROTEIN RO) (RORNP) (SJOGREN SYNDROME TYPE A ANTIGEN (SS-A)).
ACCESSION P10155
PID g133251
VERSION P10155 GI:133251
DBSOURCE swissprot: locus R060_HUMAN, accession P10155.
class: standard.
created: Mar 1, 1989.
sequence updated: Mar 1, 1989.
annotation updated: Feb 1, 1996.
xrefs: gi: gi: 177782, gi: gi: 177783, gi: gi: 387656, gi: gi: 387657, gi: gi: 86722, gi: gi: 107626
xrefs (non-sequence databases): MIM 600063, MIM 234700, PROSITE PS00030
KEYWORDS Ribonucleoprotein; RNA-binding; Systemic lupus erythematosus; Antigen.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (residues 1 to 538)
AUTHORS Deutscher, S.L., Harley, J.B. and Keene, J.D.
TITLE Molecular analysis of the 60-kDa human Ro ribonucleoprotein
JOURNAL Proc. Natl. Acad. Sci. U.S.A. 85 (24), 9479-9483 (1988)
MEDLINE 89071722
REMARK SEQUENCE FROM N.A.
REFERENCE 2 (residues 1 to 538)
AUTHORS Ben-Chetrit, E., Gandy, B.J., Tan, E.M. and Sullivan, K.F.
TITLE Isolation and characterization of a cDNA clone encoding the 60-kD component of the human SS-A/Ro ribonucleoprotein autoantigen
JOURNAL J. Clin. Invest. 83 (4), 1284-1292 (1989)
MEDLINE 89198084
REMARK SEQUENCE FROM N.A.
COMMENT
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. The original entry is available from <http://www.expasy.ch/sprot> and <http://www.ebi.ac.uk/sprot>
[FUNCTION] UNKNOWN.
[SUBUNIT] RO SMALL RIBONUCLEOPROTEINS CONSIST OF FOUR SMALL RNA MOLECULES OF 85-112 NT, EACH OF WHICH IS COMPLEXED WITH A 60 KD PROTEIN. RO RNPS MAY ALSO CONTAIN AN ADDITIONAL 52 KD PROTEIN.
[SUBCELLULAR LOCATION] CYTOPLASMIC.
[DISEASE] SERA FROM PATIENTS WITH SYSTEMIC LUPUS ERYTHEMATOSUS OFTEN CONTAIN ANTIBODIES THAT REACT WITH THE NORMAL CELLULAR RO PROTEIN AS IF THESE ANTIGEN WAS FOREIGN.
[SIMILARITY] CONTAINS 1 RNA RECOGNITION MOTIF (RNP).
[SIMILARITY] STRONG, TO XENOPUS 60 KD RO PROTEIN.
FEATURES
source Location/Qualifiers
1..538
/organism="Homo sapiens"
/db_xref="taxon:9606"
Protein 1..538
/product="60 KD RO PROTEIN"
Region 93..98
/region_name="Domain"
/note="RNA-BINDING (RNP2) (BY SIMILARITY)."
Region 124..131
/region_name="Domain"
/note="RNA-BINDING (RNP1) (BY SIMILARITY)."
Region 239
/region_name="Conflict"
/note="K -> R (IN REF. 2)."
Region 515..538
/region_name="Conflict"
/note="GMLDMCGFDTCALDVIRNPTLDMI -> ALQNTLLNKSP (IN REF. 2)."
ORIGIN
1 meesvnmqgp lnekqiansq dgyvwqvtdm nrlhrflcfc seggtypike qklglенааа
61 lirlidtdrg cevigeiksf sqegrttkqe pmlfalaics qcsdistkqa afkavsevor
121 ipthlftfiq fkkdlkesmk cgmwgralkk aiadwpnekq gmalalavtk ykqrgvshk
181 dlrlrlshlkp sseglaivtk yltkgwkev h elykekalsv etekllkyle avekvkrtd
241 elevihliee hrlvrehllt nhllkskewk allqemplita llrnlgkmta nsvlepgnse
301 vslvcekln ekllkarih pphilialet yktghglrgk lkwrpdeeil kaldaafykt
361 fktvptgkr fillavdvss mnqrvglsil nastvaaamc mvvtrtekds yvvafsdemv
421 pcvtttdmtl qqvlnamsqi paggtcdslp miwaqktntp advfivftdn etfaggvhpа
481 ialreyrkkm dipaklivcg mtsngftiad pddrgmldmc gfdtgaldvi rnftldmi
//

FIG. 22A

File Edit Search Type Help

Accession Code Query

inpharmatica

Tools

Sequences

Ligands

Select a database and enter a code:

Database	
GenBank	
Accession	<input type="radio"/>
Accession.Version	<input type="radio"/>
GI-NID	<input type="radio"/>
protein_id	<input type="radio"/>
GI-PID	<input type="radio"/>
SWISS-PROT	
Accession	<input checked="" type="radio"/>
ID	<input type="radio"/>
PDB	
ID	<input type="radio"/>

P10155

Submit Reset

RP10155: 60 KDA RO PROTEIN (60 KDA RIBONUCLEOPROTEIN RO) (RORNP) (SJORGEN SYNDROME TYPE A ANTIGEN (SS-A)).

[illegible]

FIG. 22B

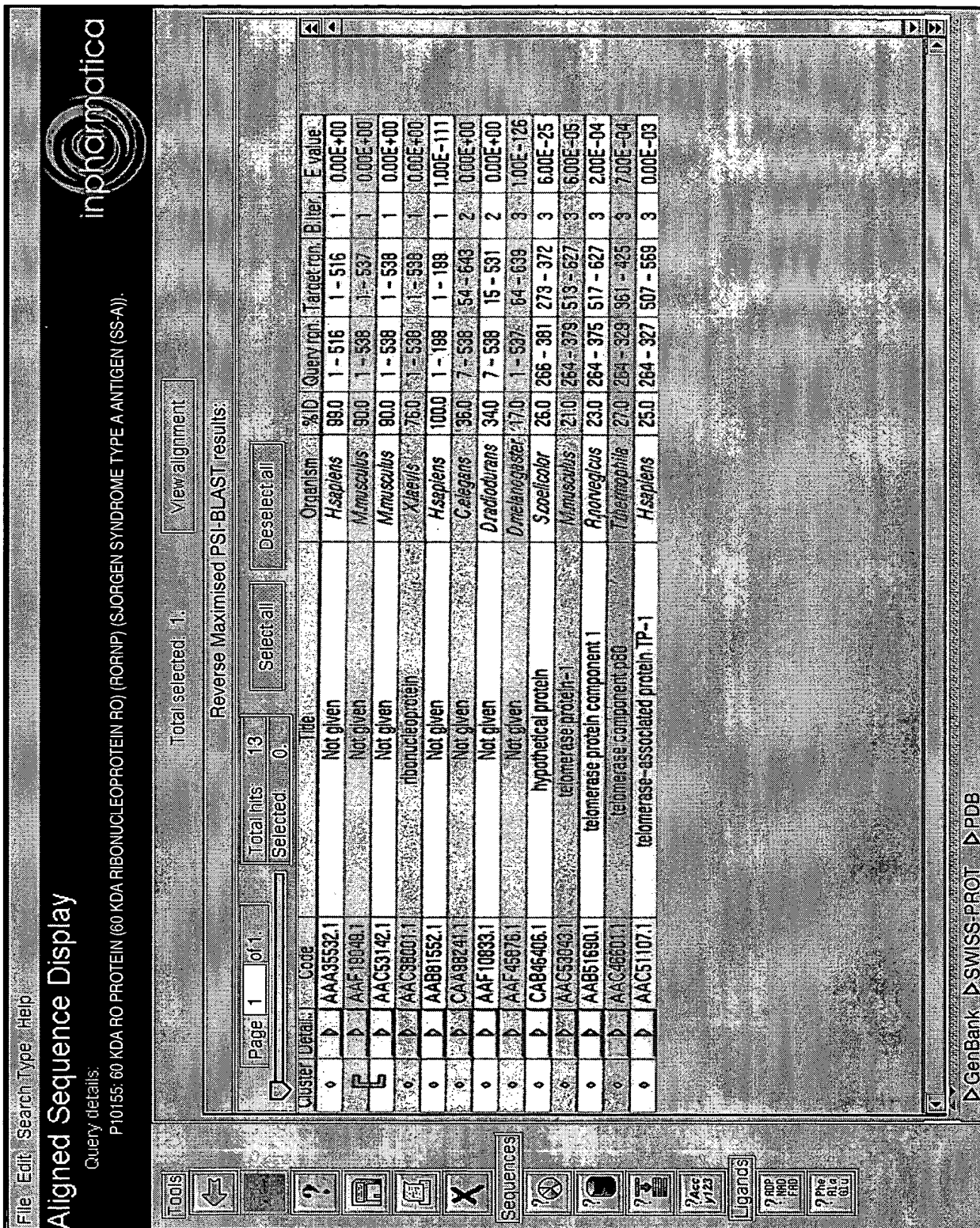


FIG. 23

AI Eye output (January 2, 2001 1:29 PM)

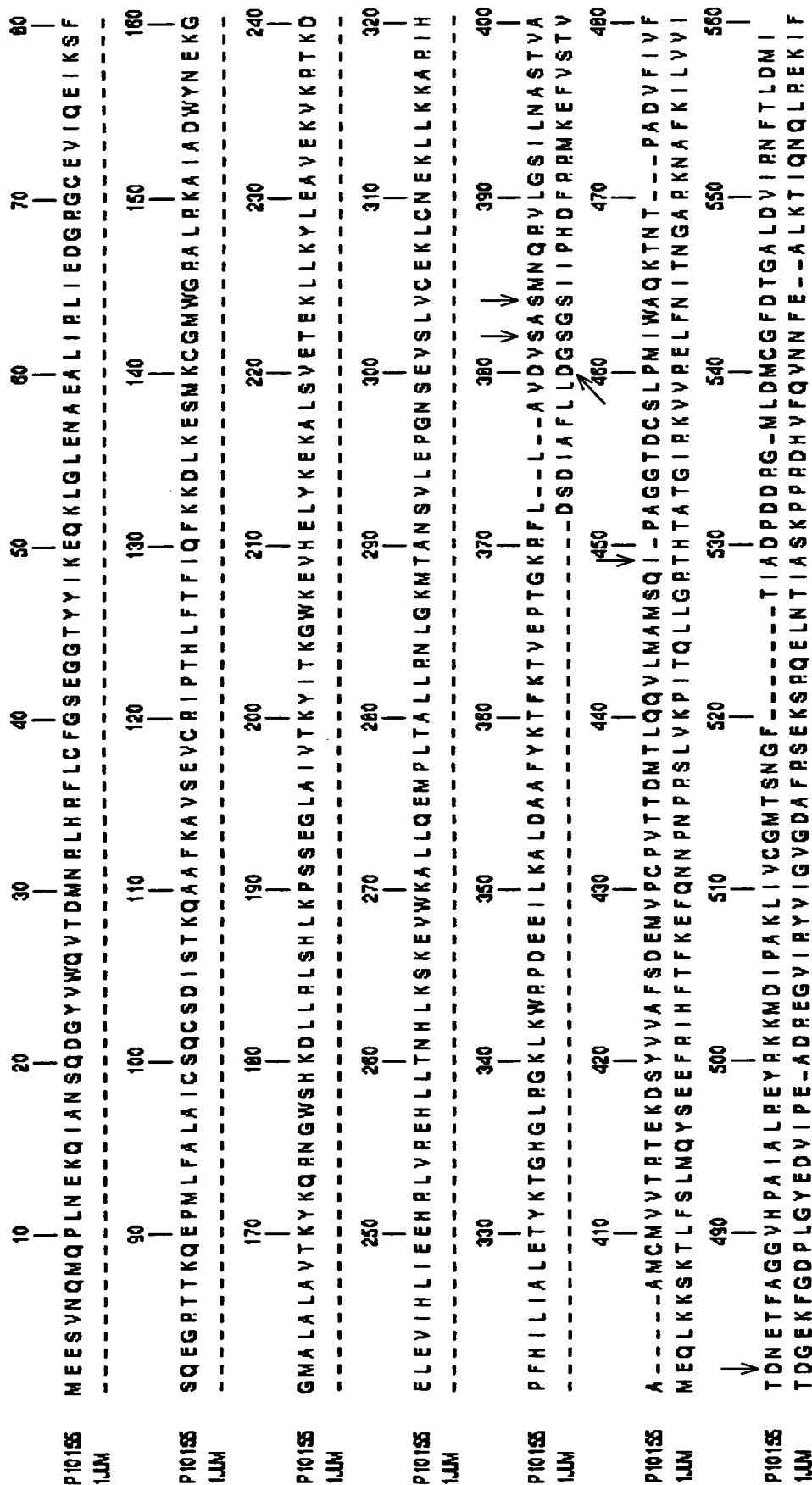
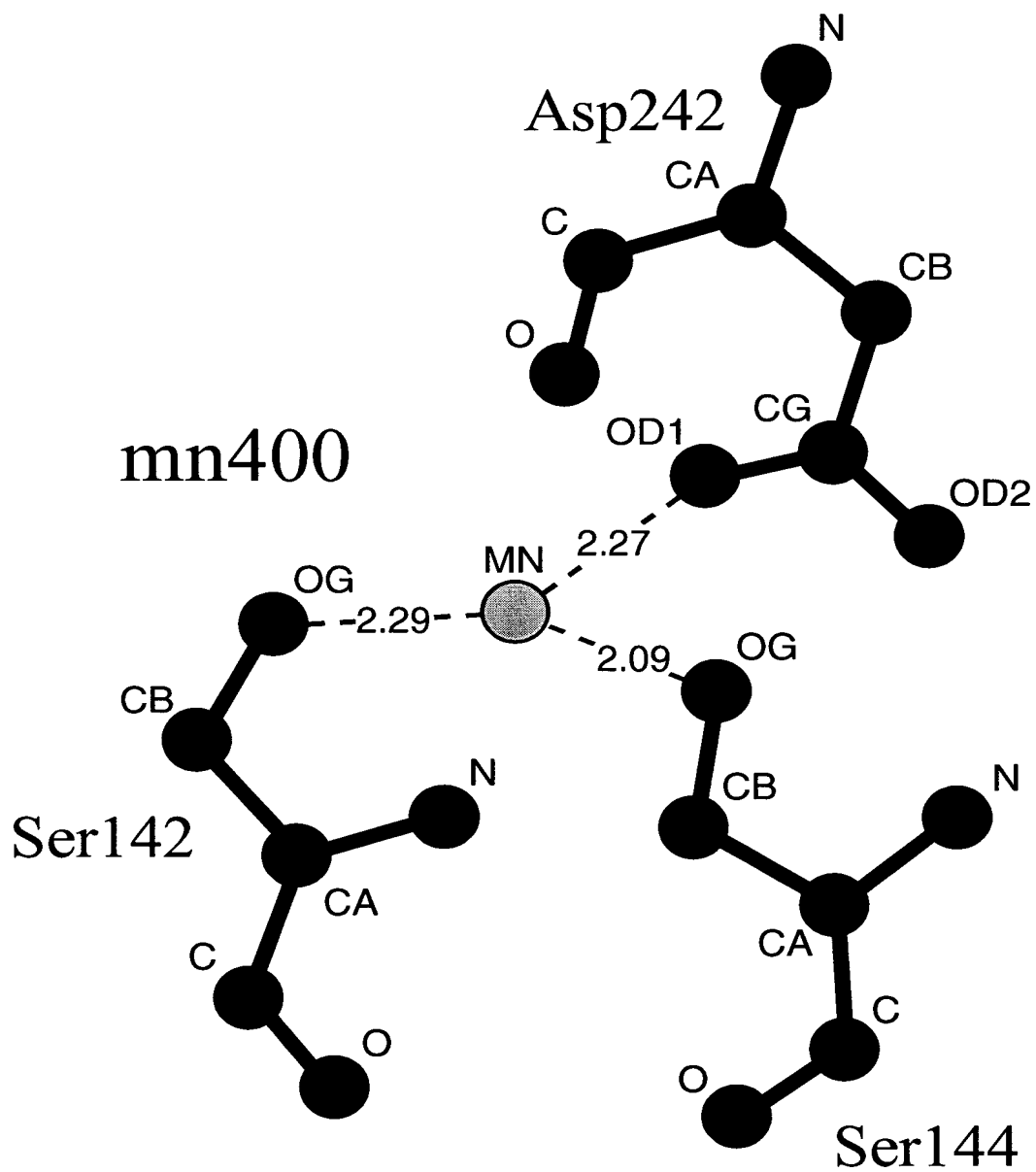


FIG. 24A



1jlm: MN 400

FIG. 24B

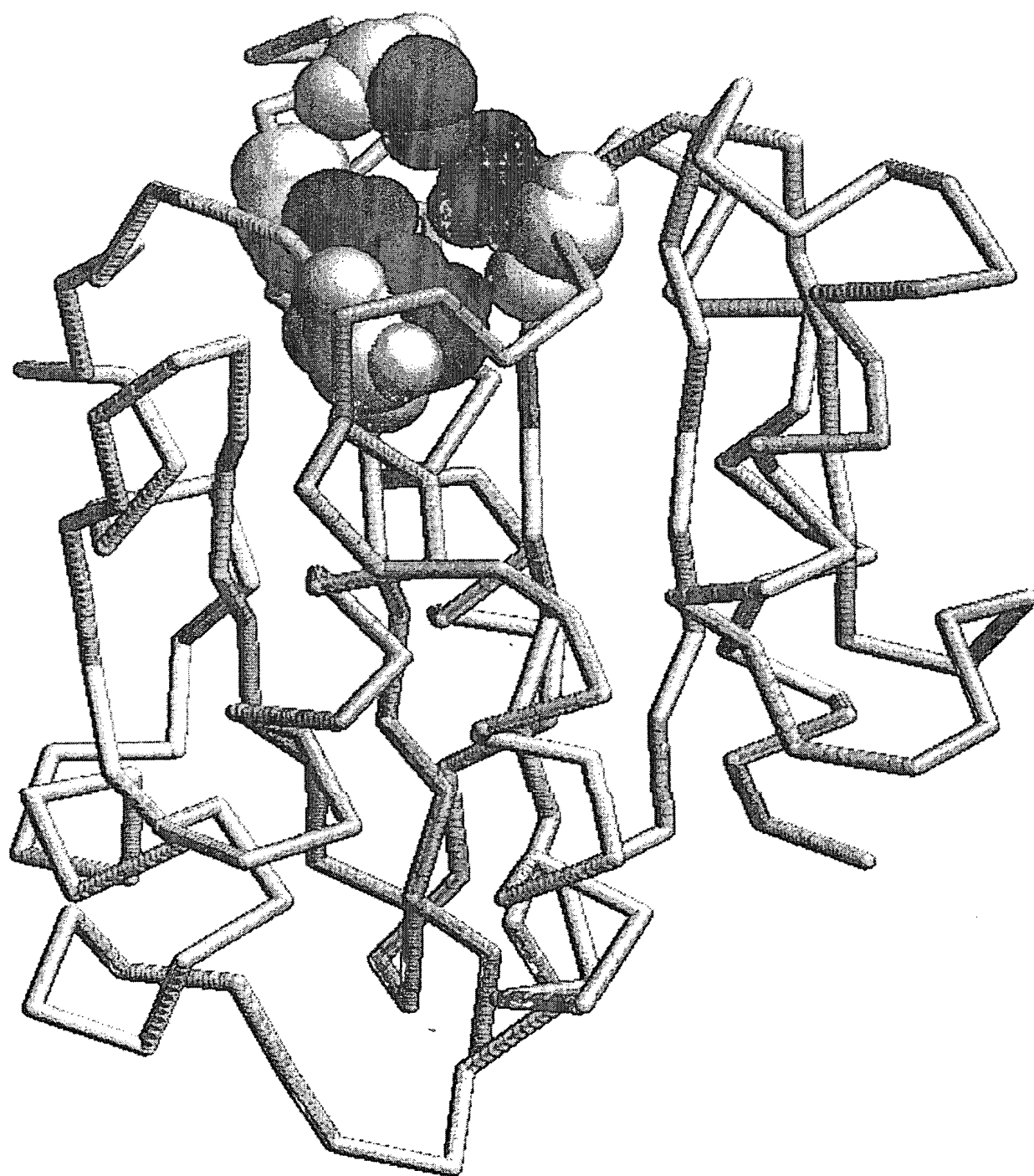


FIG. 25 (contd.)

AIEye output (January 4, 2001 3:18 PM)

470 480 490 500 510 520 530 540
↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓
P10155 YKTFKTVPTGKRFLLAVDVSA SMNQRLVLSILNASTVAAAMCMVVVTRTEKDSYVVAFSDEMVPVPTTDMTLQQVLM
AAF19049.1 YTFKTVPTGKRFLLAVDVSA SMNQRLVLSILNASTVAAAMCMVVVTRTEKESVVAFACDMVPFPVTTDMTLQQVLT
AAC38001.1 YKSFKLVEPTGKRFLLAIDVSA SMNQRLVLSILNASTVAAAMCMVVA RTEKDSHMVAFSDEMLPCPIITVNMLLHEVVE
CAA98241.1 YKAFVNA PPTGKRYCLALDVSGSMTSRVSSPLSCREAA TGMSLINLHNEAEVRCVAFCDKLTLPFTKDWKIGQVND

550 560 570 580 590 600 610 620
↓ ↓ ↓ ↓ ↓ ↓ ↓ ↓
P10155 AMSQIPAGGTDCSLPMIWAQKTNTPADVFIVFTDNETFAGGVHPAIALREYRKMDI-PAKLI VCGMTSNGFTIADPD
AAF19049.1 AMNKVPAGNTDCSLPMIWAQKTGTAAADVFI VFTDNETFAGQVHPAVALREYRKMDI-PAKLI VCGMTSNGFTIADPD
AAC38001.1 KMSDITMGSTDICALPMLWAQKTNATAADIFI VFTDCETNVEDVHPATALKQYREKMG I-PAKLI VCAMTSNGFSIADPD
CAA98241.1 YVNNLDFGRTDCGLPMTWATENNLFKDFVFI YTDNDTWAGEIHPFEAIKKYREASGIHDAKVI VMAMQAYDYSIADPS

630 640 650
↓ ↓ ↓
P10155 DRGMLDMCGFDTGALDVI RNFTLDM I
AAF19049.1 DRGMLDMCGFDTAALDVI RNFTLDVI
AAC38001.1 DRGMLDICGFD SGALDVI RNFTLDLI
CAA98241.1 DAGMLDITGFD SAVPQIVHEFVTGKI